

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/506,448A
Source: PG/10
Date Processed by STIC: 6/15/05

ENTERED



PCT

RAW SEQUENCE LISTING

DATE: 06/15/2005

PATENT APPLICATION: US/10/506,448A

TIME: 15:52:03

Input Set : A:\SEQUENCE LISTING.ST25.txt

Output Set: N:\CRF4\06152005\J506448A.raw

```

3 <110> APPLICANT: Plantechno S.r.l.
4   Plantechno S.r.l.
6 <120> TITLE OF INVENTION: IN-SEED LYSOSOMAL ENZYMES
8 <130> FILE REFERENCE: BW266R/RVP/rmp
C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/506,448A
C--> 10 <141> CURRENT FILING DATE: 2004-09-01
10 <150> PRIOR APPLICATION NUMBER: RM2002A000115
11 <151> PRIOR FILING DATE: 2002-01-03
13 <160> NUMBER OF SEQ ID NOS: 15
15 <170> SOFTWARE: PatentIn version 3.1
17 <210> SEQ ID NO: 1
18 <211> LENGTH: 1551
19 <212> TYPE: DNA
20 <213> ORGANISM: Homo sapiens
22 <220> FEATURE:
23 <221> NAME/KEY: sig_peptide
24 <222> LOCATION: (1)..(57)
25 <223> OTHER INFORMATION:
27 <220> FEATURE:
28 <221> NAME/KEY: CDS
29 <222> LOCATION: (1)..(1548)
30 <223> OTHER INFORMATION:
W--> 32 <220>
33 <221> NAME/KEY: mat_peptide
34 <222> LOCATION: (58)..()
35 <223> OTHER INFORMATION:
W--> 37 <400> 1
38 atg gct ggc agc ctc aca ggt ttg ctt cta ctt cag gca gtg tcg tgg      48
39 Met Ala Gly Ser Leu Thr Gly Leu Leu Leu Leu Gln Ala Val Ser Trp
40           -15          -10          -5
42 gca tca ggt gcc cgc ccc tgc atc cct aaa agc ttc ggc tac agc tcg      96
43 Ala Ser Gly Ala Arg Pro Cys Ile Pro Lys Ser Phe Gly Tyr Ser Ser
44           -1  1          5          10
46 gtg gtg tgt gtc tgc aat gcc aca tac tgt gac tcc ttt gac ccc ccg      144
47 Val Val Cys Val Cys Asn Ala Thr Tyr Cys Asp Ser Phe Asp Pro Pro
48    15    20    25
50 acc ttt cct gcc ctt ggt acc ttc agc cgc tat gag agt aca cgc agt      192
51 Thr Phe Pro Ala Leu Gly Thr Phe Ser Arg Tyr Glu Ser Thr Arg Ser
52 30          35          40          45
54 ggg cga cgg atg gag ctg agt atg ggg ccc atc cag gct aat cac acg      240
55 Gly Arg Arg Met Glu Leu Ser Met Gly Pro Ile Gln Ala Asn His Thr
56          50          55          60
58 ggc aca ggc ctg cta ctg acc ctg cag cca gaa cag aag ttc cag aaa      288

```

RAW SEQUENCE LISTING

DATE: 06/15/2005

PATENT APPLICATION: US/10/506,448A

TIME: 15:52:03

Input Set : A:\SEQUENCE LISTING.ST25.txt

Output Set: N:\CRF4\06152005\J506448A.raw

59	Gly	Thr	Gly	Leu	Leu	Leu	Thr	Leu	Gln	Pro	Glu	Gln	Lys	Phe	Gln	Lys	
60				65				70					75				
62	gtg	aag	gga	ttt	gga	ggg	gcc	atg	aca	gat	gct	gct	gct	ctc	aac	atc	336
63	Val	Lys	Gly	Phe	Gly	Gly	Ala	Met	Thr	Asp	Ala	Ala	Ala	Leu	Asn	Ile	
64			80					85					90				
66	ctt	gcc	ctg	tca	ccc	cct	gcc	caa	aat	ttg	cta	ctt	aaa	tgc	tac	ttc	384
67	Leu	Ala	Leu	Ser	Pro	Pro	Ala	Gln	Asn	Leu	Leu	Leu	Lys	Ser	Tyr	Phe	
68		95					100					105					
70	tct	gaa	gaa	gga	atc	gga	tat	aac	atc	atc	cgg	gta	ccc	atg	gcc	agc	432
71	Ser	Glu	Glu	Gly	Ile	Gly	Tyr	Asn	Ile	Ile	Arg	Val	Pro	Met	Ala	Ser	
72	110					115					120					125	
74	tgt	gac	ttc	tcc	atc	cgc	acc	tac	acc	tat	gca	gac	acc	cct	gat	gat	480
75	Cys	Asp	Phe	Ser	Ile	Arg	Thr	Tyr	Thr	Tyr	Ala	Asp	Thr	Pro	Asp	Asp	
76				130						135					140		
78	ttc	cag	ttg	cac	aac	ttc	agc	ctc	cca	gag	gaa	gat	acc	aag	ctc	aag	528
79	Phe	Gln	Leu	His	Asn	Phe	Ser	Leu	Pro	Glu	Glu	Asp	Thr	Lys	Leu	Lys	
80				145					150					155			
82	ata	ccc	ctg	att	cac	cga	gcc	ctg	cag	ttg	gcc	cag	cgt	ccc	gtt	tca	576
83	Ile	Pro	Leu	Ile	His	Arg	Ala	Leu	Gln	Leu	Ala	Gln	Arg	Pro	Val	Ser	
84			160					165					170				
86	ctc	ctt	gcc	agc	ccc	tgg	aca	tca	ccc	act	tgg	ctc	aag	acc	aat	gga	624
87	Leu	Leu	Ala	Ser	Pro	Trp	Thr	Ser	Pro	Thr	Trp	Leu	Lys	Thr	Asn	Gly	
88		175					180					185					
90	gcg	gtg	aat	ggg	aag	ggg	tca	ctc	aag	gga	cag	ccc	gga	gac	atc	tac	672
91	Ala	Val	Asn	Gly	Lys	Gly	Ser	Leu	Lys	Gly	Gln	Pro	Gly	Asp	Ile	Tyr	
92	190					195					200					205	
94	cac	cag	acc	tgg	gcc	aga	tac	ttt	gtg	aag	ttc	ctg	gat	gcc	tat	gct	720
95	His	Gln	Thr	Trp	Ala	Arg	Tyr	Phe	Val	Lys	Phe	Leu	Asp	Ala	Tyr	Ala	
96				210						215					220		
98	gag	cac	aag	tta	cag	ttc	tgg	gca	gtg	aca	gct	gaa	aat	gag	cct	tct	768
99	Glu	His	Lys	Leu	Gln	Phe	Trp	Ala	Val	Thr	Ala	Glu	Asn	Glu	Pro	Ser	
100				225					230					235			
102	gct	ggg	ctg	ttg	agt	gga	tac	ccc	ttc	cag	tgc	ctg	ggc	ttc	acc	cct	816
103	Ala	Gly	Leu	Leu	Ser	Gly	Tyr	Pro	Phe	Gln	Cys	Leu	Gly	Phe	Thr	Pro	
104			240					245					250				
106	gaa	cat	cag	cga	gac	ttc	att	gcc	cgt	gac	cta	ggt	cct	acc	ctc	gcc	864
107	Glu	His	Gln	Arg	Asp	Phe	Ile	Ala	Arg	Asp	Leu	Gly	Pro	Thr	Leu	Ala	
108		255					260					265					
110	aac	agt	act	cac	cac	aat	gtc	cgc	cta	ctc	atg	ctg	gat	gac	caa	cgc	912
111	Asn	Ser	Thr	His	His	Asn	Val	Arg	Leu	Leu	Met	Leu	Asp	Asp	Gln	Arg	
112	270					275					280					285	
114	ttg	ctg	ctg	ccc	cac	tgg	gca	aag	gtg	gta	ctg	aca	gac	cca	gaa	gca	960
115	Leu	Leu	Leu	Pro	His	Trp	Ala	Lys	Val	Val	Leu	Thr	Asp	Pro	Glu	Ala	
116				290						295					300		
118	gct	aaa	tat	gtt	cat	ggc	att	gct	gta	cat	tgg	tac	ctg	gac	ttt	ctg	1008
119	Ala	Lys	Tyr	Val	His	Gly	Ile	Ala	Val	His	Trp	Tyr	Leu	Asp	Phe	Leu	
120			305					310					315				
122	gct	cca	gcc	aaa	gcc	acc	cta	ggg	gag	aca	cac	cgc	ctg	ttc	ccc	aac	1056
123	Ala	Pro	Ala	Lys	Ala	Thr	Leu	Gly	Glu	Thr	His	Arg	Leu	Phe	Pro	Asn	

RAW SEQUENCE LISTING

DATE: 06/15/2005

PATENT APPLICATION: US/10/506,448A

TIME: 15:52:03

Input Set : A:\SEQUENCE LISTING.ST25.txt

Output Set: N:\CRF4\06152005\J506448A.raw

```

124          320          325          330
126 acc atg ctc ttt gcc tca gag gcc tgt gtg ggc tcc aag ttc tgg gag      1104
127 Thr Met Leu Phe Ala Ser Glu Ala Cys Val Gly Ser Lys Phe Trp Glu
128          335          340          345
130 cag agt gtg cgg cta ggc tcc tgg gat cga ggg atg cag tac agc cac      1152
131 Gln Ser Val Arg Leu Gly Ser Trp Asp Arg Gly Met Gln Tyr Ser His
132 350          355          360          365
134 agc atc atc acg aac ctc ctg tac cat gtg gtc ggc tgg acc gac tgg      1200
135 Ser Ile Ile Thr Asn Leu Leu Tyr His Val Val Gly Trp Thr Asp Trp
136          370          375          380
138 aac ctt gcc ctg aac ccc gaa gga gga ccc aat tgg gtg cgt aac ttt      1248
139 Asn Leu Ala Leu Asn Pro Glu Gly Gly Pro Asn Trp Val Arg Asn Phe
140          385          390          395
142 gtc gac agt ccc atc att gta gac atc acc aag gac acg ttt tac aaa      1296
143 Val Asp Ser Pro Ile Ile Val Asp Ile Thr Lys Asp Thr Phe Tyr Lys
144          400          405          410
146 cag ccc atg ttc tac cac ctt ggc cac ttc agc aag ttc att cct gag      1344
147 Gln Pro Met Phe Tyr His Leu Gly His Phe Ser Lys Phe Ile Pro Glu
148          415          420          425
150 ggc tcc cag aga gtg ggg ctg gtt gcc agt cag aag aac gac ctg gac      1392
151 Gly Ser Gln Arg Val Gly Leu Val Ala Ser Gln Lys Asn Asp Leu Asp
152 430          435          440          445
154 gca gtg gca ctg atg cat ccc gat ggc tct gct gtt gtg gtc gtg cta      1440
155 Ala Val Ala Leu Met His Pro Asp Gly Ser Ala Val Val Val Val Leu
156          450          455          460
158 aac cgc tcc tct aag gat gtg cct ctt acc atc aag gat cct gct gtg      1488
159 Asn Arg Ser Ser Lys Asp Val Pro Leu Thr Ile Lys Asp Pro Ala Val
160          465          470          475
162 ggc ttc ctg gag aca atc tca cct ggc tac tcc att cac acc tac ctg      1536
163 Gly Phe Leu Glu Thr Ile Ser Pro Gly Tyr Ser Ile His Thr Tyr Leu
164          480          485          490
166 tgg cat cgc cag tga      1551
167 Trp His Arg Gln
168          495
171 <210> SEQ ID NO: 2
172 <211> LENGTH: 516
173 <212> TYPE: PRT
174 <213> ORGANISM: Homo sapiens
176 <400> SEQUENCE: 2
178 Met Ala Gly Ser Leu Thr Gly Leu Leu Leu Leu Gln Ala Val Ser Trp
179          -15          -10          -5
182 Ala Ser Gly Ala Arg Pro Cys Ile Pro Lys Ser Phe Gly Tyr Ser Ser
183          -1  1          5          10
186 Val Val Cys Val Cys Asn Ala Thr Tyr Cys Asp Ser Phe Asp Pro Pro
187          15          20          25
190 Thr Phe Pro Ala Leu Gly Thr Phe Ser Arg Tyr Glu Ser Thr Arg Ser
191 30          35          40          45
194 Gly Arg Arg Met Glu Leu Ser Met Gly Pro Ile Gln Ala Asn His Thr
195          50          55          60

```

RAW SEQUENCE LISTING

DATE: 06/15/2005

PATENT APPLICATION: US/10/506,448A

TIME: 15:52:03

Input Set : A:\SEQUENCE LISTING.ST25.txt

Output Set: N:\CRF4\06152005\J506448A.raw

```

198 Gly Thr Gly Leu Leu Leu Thr Leu Gln Pro Glu Gln Lys Phe Gln Lys
199      65      70      75
202 Val Lys Gly Phe Gly Gly Ala Met Thr Asp Ala Ala Ala Leu Asn Ile
203      80      85      90
206 Leu Ala Leu Ser Pro Pro Ala Gln Asn Leu Leu Leu Lys Ser Tyr Phe
207      95     100     105
210 Ser Glu Glu Gly Ile Gly Tyr Asn Ile Ile Arg Val Pro Met Ala Ser
211 110     115     120     125
214 Cys Asp Phe Ser Ile Arg Thr Tyr Thr Tyr Ala Asp Thr Pro Asp Asp
215     130     135     140
218 Phe Gln Leu His Asn Phe Ser Leu Pro Glu Glu Asp Thr Lys Leu Lys
219     145     150     155
222 Ile Pro Leu Ile His Arg Ala Leu Gln Leu Ala Gln Arg Pro Val Ser
223     160     165     170
226 Leu Leu Ala Ser Pro Trp Thr Ser Pro Thr Trp Leu Lys Thr Asn Gly
227     175     180     185
230 Ala Val Asn Gly Lys Gly Ser Leu Lys Gly Gln Pro Gly Asp Ile Tyr
231 190     195     200     205
234 His Gln Thr Trp Ala Arg Tyr Phe Val Lys Phe Leu Asp Ala Tyr Ala
235     210     215     220
238 Glu His Lys Leu Gln Phe Trp Ala Val Thr Ala Glu Asn Glu Pro Ser
239     225     230     235
242 Ala Gly Leu Leu Ser Gly Tyr Pro Phe Gln Cys Leu Gly Phe Thr Pro
243     240     245     250
246 Glu His Gln Arg Asp Phe Ile Ala Arg Asp Leu Gly Pro Thr Leu Ala
247     255     260     265
250 Asn Ser Thr His His Asn Val Arg Leu Leu Met Leu Asp Asp Gln Arg
251 270     275     280     285
254 Leu Leu Leu Pro His Trp Ala Lys Val Val Leu Thr Asp Pro Glu Ala
255     290     295     300
258 Ala Lys Tyr Val His Gly Ile Ala Val His Trp Tyr Leu Asp Phe Leu
259     305     310     315
262 Ala Pro Ala Lys Ala Thr Leu Gly Glu Thr His Arg Leu Phe Pro Asn
263     320     325     330
266 Thr Met Leu Phe Ala Ser Glu Ala Cys Val Gly Ser Lys Phe Trp Glu
267     335     340     345
270 Gln Ser Val Arg Leu Gly Ser Trp Asp Arg Gly Met Gln Tyr Ser His
271 350     355     360     365
274 Ser Ile Ile Thr Asn Leu Leu Tyr His Val Val Gly Trp Thr Asp Trp
275     370     375     380
278 Asn Leu Ala Leu Asn Pro Glu Gly Gly Pro Asn Trp Val Arg Asn Phe
279     385     390     395
282 Val Asp Ser Pro Ile Ile Val Asp Ile Thr Lys Asp Thr Phe Tyr Lys
283     400     405     410
286 Gln Pro Met Phe Tyr His Leu Gly His Phe Ser Lys Phe Ile Pro Glu
287     415     420     425
290 Gly Ser Gln Arg Val Gly Leu Val Ala Ser Gln Lys Asn Asp Leu Asp
291 430     435     440     445
294 Ala Val Ala Leu Met His Pro Asp Gly Ser Ala Val Val Val Val Leu

```

RAW SEQUENCE LISTING

DATE: 06/15/2005

PATENT APPLICATION: US/10/506,448A

TIME: 15:52:03

Input Set : A:\SEQUENCE LISTING.ST25.txt

Output Set: N:\CRF4\06152005\J506448A.raw

```

295           450           455           460
298 Asn Arg Ser Ser Lys Asp Val Pro Leu Thr Ile Lys Asp Pro Ala Val
299           465           470           475
302 Gly Phe Leu Glu Thr Ile Ser Pro Gly Tyr Ser Ile His Thr Tyr Leu
303           480           485           490
306 Trp His Arg Gln
307           495
310 <210> SEQ ID NO: 3
311 <211> LENGTH: 27
312 <212> TYPE: DNA
313 <213> ORGANISM: Artificial Sequence
315 <220> FEATURE:
316 <223> OTHER INFORMATION: forward primer for human GCB amplification including DNA
coding f
317           or native signal peptide
319 <400> SEQUENCE: 3
320 tctagaatgg ctggcagcct cacaggt                               27
323 <210> SEQ ID NO: 4
324 <211> LENGTH: 33
325 <212> TYPE: DNA
326 <213> ORGANISM: Artificial Sequence
328 <220> FEATURE:
329 <223> OTHER INFORMATION: Reverse primer for human GCB amplification
331 <400> SEQUENCE: 4
332 gtgtggatgg acaccgtagc ggtcactctc gag                               33
335 <210> SEQ ID NO: 5
336 <211> LENGTH: 31
337 <212> TYPE: DNA
338 <213> ORGANISM: Artificial Sequence
340 <220> FEATURE:
341 <223> OTHER INFORMATION: forward primer for human GCB amplification excluding DNA
coding f
342           or native signal peptide
344 <400> SEQUENCE: 5
345 cccgggtgcc cgcccctgca tccctaaaag c                               31
348 <210> SEQ ID NO: 6
349 <211> LENGTH: 1428
350 <212> TYPE: DNA
351 <213> ORGANISM: Glycine max
353 <220> FEATURE:
354 <221> NAME/KEY: promoter
355 <222> LOCATION: (1)..(1428)
356 <223> OTHER INFORMATION:
W--> 358 <400> 6
359 taaaataatc tatacattaa aaaatttgat tttaaattt tagaaattca tgattttatt       60
361 tttttttacc agaaatccgt taatattggt aaaatattac caactaattt ataaatttta       120
363 ttttaaggca attaagcatg tttgataaaa tatatatatt gttataaata cttttcaaaa       180
365 gtataaagtt gatgatggcg tgggtggtaga ttattttagt tctaggttcg aatgcaagtt       240
367 gggtttagaca tttagcctta ttcttttttc taacccaaat aaatgtaaat ggaaaacctt       300
369 taggaaaaaa aagaaatcaa aattgaaaac atcatccggt ggagtcgaga agccacacc       360
371 cacgtgaccc aacaatatta aaataagagt ttgctctaca gtaaatgcga tactttttta       420

```

VERIFICATION SUMMARY

DATE: 06/15/2005

PATENT APPLICATION: US/10/506,448A

TIME: 15:52:04

Input Set : A:\SEQUENCE LISTING.ST25.txt

Output Set: N:\CRF4\06152005\J506448A.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application No
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:32 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:25
L:37 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:30
L:37 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:35
L:358 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:6,Line#:356
L:418 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:7,Line#:416
L:439 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:8,Line#:432
L:444 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:8,Line#:437
L:444 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:8,Line#:442
L:713 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:12,Line#:706
L:718 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:12,Line#:711
L:718 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:12,Line#:716